Innovation and Technology Studies
ISSN: 3007-6927
DOI: https://doi.org/10.61784/its3009

MORAL DECISION-MAKING FOR AUTONOMOUS DRIVING BASED ON MULTI-OBJECTIVE REINFORCEMENT LEARNING

XiangYu Chen
College of Social Sciences, Shenzhen University, Shenzhen 518060, Guangdong, China.
Corresponding Email: cxybeing@gmail.com

Abstract: Vehicles must ensure safety and efficiency and deal with complex ethical dilemmas in autonomous driving. In order to deal with these ethical dilemmas effectively, moral decision-making models based on multi-objective reinforcement learning (MORL) provide a technical path to resolve such ethical dilemmas. Unlike traditional reinforcement learning (RL), MORL can generate more socially moral decision-making strategies in conflict scenarios by simultaneously optimizing multiple objectives. Of course, significant challenges remain in this research path. Assigning reward function weights is highly dependent on subjective judgement and cultural context; the dynamic environment is not adaptable enough, and the scarcity of ethical dilemma data limits model training. To address these issues, this paper points out that future research needs to focus on the dynamic weight adjustment mechanism, the construction of cross-cultural ethical frameworks, and large-scale real-world validation.
Keywords: Autonomous driving; Moral decision-making; Ethical dilemmas; Multi-objective reinforcement learning

1 INTRODUCTION

Autonomous driving is gradually moving from the laboratory into real life. However, how to make autonomous driving systems make moral decisions that meet ethical standards in complex traffic environments has become a difficult problem in current technology practice [1-4]. Commonly used in traditional intelligent decision-making systems is the RL decision-making framework, which aims to achieve single-objective optimization by prompting the AI agent to take the best action amongst reward-maximizing and corresponding constraint strategies [5-8]. Although traditional RL excels at single-objective optimization (e.g., minimizing collision rates or maximizing traffic efficiency), its inherent flaws are exposed in ethical multi-objective conflict scenarios. To address this challenge, MORL provides a new path to solving ethical dilemmas by simultaneously optimizing conflicting objectives and a dynamic weight allocation mechanism. Unlike RL's single reward function, MORL allows for the definition of multi-dimensional objective functions (e.g., safety, efficiency, fairness) and identifies the optimal trade-off solution via the Pareto front. This ability makes MORL significantly superior to traditional methods regarding cross-cultural ethical adaptability. Researchers have attempted to incorporate multiple ethical objectives in intelligent decision-making systems to achieve decision optimization that is more in line with socio-ethical rules [9-12].

2 MORAL DECISION-MAKING FOR AUTONOMOUS DRIVING

2.1 Dilemma of Moral Decision-Making in Autonomous Driving

The Trolley Problem (TP), a classic thought experiment in ethics, explores how one should make decisions in ethical dilemmas to maximize benefits when faced with unavoidable harm [13-15]. For autonomous driving decision-making systems, this dilemma manifests itself in how to make appropriate choices to achieve optimal ethical goals in inevitable traffic accidents [16-20]. Autonomous driving is often faced with a situation where they must choose between hitting a pedestrian, hitting another vehicle, or some other more damaging scenario to minimize harm (See figure 1). In such situations, it is complex to make an accurate judgement based solely on a simple utilitarian morality (although this is currently the most adopted ethical rule in moral decision-making), which requires a more complex and flexible moral decision-making framework to assist autonomous driving in making their choices.

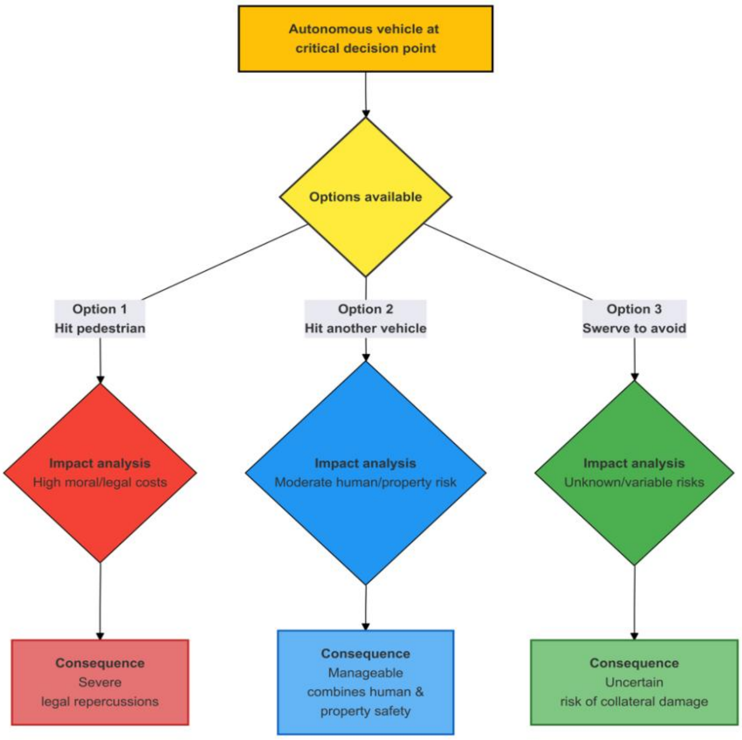

Figure 1 Autonomous vehicle at critical decision point

Of course, some scholars are pessimistic about the excessive focus on TP in ethical discussions of autonomous driving. Etienne argues that current approaches to AI ethics in the autonomous driving industry often simplify complex moral dilemmas and ignore broader social, cultural and situational factors [21]. Although frameworks such as the TP are widely discussed, they fail to effectively address the systemic risks and operational challenges that autonomous driving faces in real-world environments. Meanwhile, Geisslinger et al. think ethical considerations of autonomous driving should go beyond hypothetical moral dilemmas and focus on practical situations involving risk management, safety, and societal impacts [2]. Tolmeijer et al. argue that while AI decision-making systems can assist in moral decision-making by providing objective analyses, identifying patterns, or exploring options, they are inherently non-moral decision-making and still rely on human experts to provide moral reasoning and judgement [22]. AI moral decision-making should be a tool to support humans in achieving moral judgement rather than a subject (human being) who substitutes for moral reasoning. Addressing the issue of explainability in AI moral decision-making, Madhav and Tyagi point to explainable artificial intelligence (XAI) as the key to solving the problem of lack of trust in autonomous driving [23]. By making AI decisions transparent, explainable and consistent with human expectations, XAI can enhance trust, drive the adoption of autonomous driving technologies and ensure their safe and ethical integration into society. Despite the current growing concern in the autonomous driving car industry about ethical issues such as safety, liability, bias in decision-making algorithms, and social acceptance, there is still a lack of consensus on how to comprehensively address these issues [24,25]. This requires interdisciplinary collaboration, more transparent regulations, and the alignment of technological developments with societal values to ensure the ethical deployment of autonomous driving.

2.2 Frameworks for Moral Decision-Making in Autonomous Driving

An AI moral decision-making framework refers to some systematic mechanism and model that ensures that the decisions made by an intelligent decision-making system are in line with ethical principles and societal expectations when faced with complex situations. Conitzer et al. argue that moral decision-making frameworks typically rely on datasets (e.g., surveys, experiments) collected from human responses to ethical dilemmas, which form the basis for training and evaluating AI models [26]. Noothigattu et al. advocate a democratic, collective approach - by aggregating the ethical preferences of a population - to allow AI moral decision-making to be based on the collective will [27]. That is, modelling moral decision-making as a voting problem, where individuals are free to express their preferences for different ethical choices, and then aggregating preferences through voting theory to ultimately select the decision consistent with the majority's ethical outlook. Another empirical study similarly suggests that consumer acceptance of self-driving car services is driven by technological advances and significantly influenced by the perceived consistency of ethical practices and personal values [28]. Regardless of the moral decision-making framework adopted for autonomous driving systems, the core components and principles remain the same (See Table 1), requiring trade-offs between multiple goals such as safety, efficiency, and the rights of passengers and pedestrians, which is the core pathway to achieving trust and legitimization of autonomous driving systems.

Table 1 The content and principles of the moral decision-making framework for autonomous driving

Content		Principle	
Multi-objective optimization	Need to trade-off between multiple objectives (e.g. traffic safety, passenger experience, pedestrian protection)	Damage minimization	Harm minimization is the most fundamental ethical requirement in moral decision-making.
Transparency	The decision-making process of an automated driving system should be transparent so that it is easy for humans to understand and trust.	Fairness and justice	In a moral decision-making framework, the system must ensure that harm is minimized and that decisions are fair and just.
Consistency	The system's decisions should demonstrate consistency in the same or similar situations to enhance public trust in self-driving vehicles.	Legal compliance	Autonomous driving systems should follow current traffic rules as much as possible, choosing the option that minimizes liability when violations and accidents are unavoidable.
Operability	The moral decision-making of the system needs to be actionable, i.e., it can be implemented algorithmically.	Interpretability	In an accident or complication with an automated driving system, the system should be able to explain clearly why it has acted in a particular way.

The moral decision-making framework for autonomous driving aims to guide vehicles in making ethical and socially desirable decisions in complex situations. The core issues include dealing with ethical dilemmas, weighing different ethical principles (e.g., utilitarianism vs. deontology), ensuring fairness and transparency, and satisfying legal and public expectations. Three broad categories of moral decision-making frameworks exist, namely rule-based, utility-maximization and social contract-based frameworks (see Table 2), each with different strengths and weaknesses in dealing with specific ethical situations.

Table 2 Advantages and Disadvantages of Different Moral Decision-Making Frameworks for Autonomous Driving

Type	Advantage	Disadvantage
Rule Based	● Realisability and interpretability ● The strong binding nature of laws and social norms	● Lack of flexibility to adapt to complex and uncertain situations ● Inability to effectively prioritize different objectives when faced with multiple conflicting objectives
Utility-Based	● Ability to deal effectively with complex ethical situations and make optimal decisions by optimizing objective functions ● Greater flexibility to adjust the weighting of different objectives according to actual situations	● There is a high degree of subjectivity in the definition of utility and the assignment of weights ● In emergencies, it may not be possible to fully consider ethical and legal responsibilities, focusing only on utility-maximizing outcomes.
Social Contract Based	● Enables greater social acceptance and cultural adaptability of decision-making ● Emphasis on the cultural context of moral decision-making to ensure that the system meets local ethical standards	● In multicultural, multi-value societies, the challenge of inconsistent standards may be faced ● The social contract needs to be modelled and quantified, which is technically challenging to achieve for the time being

3 MORL IN AUTONOMOUS DRIVING

3.1 Disadvantages of RL in Autonomous Driving

RL is a machine-learning method for learning optimal strategies through environmental interaction [29-33]. The core of RL lies in adapting the behavioural strategies of intelligence through reward feedback and generally only deals with problems with a single objective, such as maximizing long-term rewards and minimizing traffic accidents. In the field of autonomous driving, RL models need a large amount of training data to approximate the optimal policy gradually; however, obtaining real-world moral dilemma data is not only costly but also has a large amount of uncertainty, which makes RL insufficient in this specific case [34,35]. In addition, to train an autonomous driving system on how to make rational decisions in different cultures and contexts, the model needs to collect a large amount of behavioural data, including pedestrians, passengers, traffic signals and other vehicles. As these data cover ethical standards and laws and regulations across multiple domains and regions, it is important to establish a clear framework for data ownership, access, and control to ensure ethical and fair outcomes [36,37]. Finally, real-world ethical dilemmas tend to be scarcity events, which are rare in naturally occurring traffic environments. However, these scarcity events have important implications for moral decision-making in autonomous driving. For example, a vehicle that suddenly encounters a pedestrian crossing the road may need to make a decision that involves a life trade-off, such situations do not occur frequently in regular driving, but when they do, the accuracy and ethics of the decision are critical. Due to the limited data for such situations, RL models often do not gain enough experience to learn effective strategies.

3.2 MORL in Autonomous Driving

The core idea of MORL is to consider multiple different objectives simultaneously in the decision-making process rather than just pursuing a single objective. Each objective has a corresponding reward function, which is usually combined into a composite objective function by weighting to guide the intelligence to make a decision. One of the core issues of MORL applied to autonomous driving is how to design the reward function and balance the individual objectives. The following are two common algorithmic frameworks for MORL and their applications in autonomous driving.

3.2.1 Weighted sum method: combining multiple objectives

The weighted sum method is one of the simplest and most used methods in MORL. This method sums the reward functions of multiple objectives with certain weights to form a composite objective function. In autonomous driving, the objectives may include 'minimize injuries', 'comfort' and 'improve driving efficiency'. The learning process is guided by assigning a weight to each objective (e.g. the weight for injury minimization can be adjusted to the highest) and then weighting and summing these objectives to form a single objective function. Its mathematical expression is typically:

$$f(x) = \sum_{i=1}^{n} \omega_i f_i(x) \tag{1}$$

Where $f(x)$ is the composite objective function which represents the weighted sum of all the objectives. $f_i(x)$ is the -th objective function, which represents the performance or performance of the system on the -th objective. ω_i is the weight associated with the -th objective function, which usually satisfies $\omega_i \geq 0$ and $\sum_{i=1}^{n} \omega_i = 1$. It is assumed that the current autonomous driving has three objectives that need to be satisfied simultaneously:

① 'Minimising damage' = $f_1(x)$;
② 'Comfort' = $f_2(x)$;
③ 'Improving driving efficiency' = $f_3(x)$.

The weights are $\omega_1, \omega_2, \omega_3$, then the expression for the weighted sum method is:

$$f(x) = \omega_1 f_1(x) + \omega_2 f_2(x) + \omega_3 f_3(x) \tag{2}$$

Of course, how to choose the appropriate weights is crucial in practice. If the weights are not set reasonably, it may lead to a specific objective being over-optimized, which ultimately leads to an imbalance in the morality of decision-making (e.g., over-emphasis on efficiency may lead to neglecting the safety of pedestrians). Therefore, a reasonable weight adjustment mechanism is crucial for successfully applying MORL in autonomous driving. To cope with the dilemma of conflicting multi-objective weights in MORL, the Pareto optimality method is an effective means of dealing with trade-offs and conflicts between multiple objectives and is particularly suitable for optimizing conflicting objectives [38,39]. The method aims to find a compromise solution that makes it impossible to further optimize an objective without compromising other objectives, and these compromise solutions are called Pareto optimal solutions, denoted as:

$$\mathbf{x}^* \text{ is Pareto optimal } \Leftrightarrow \nexists \mathbf{x} \in \boldsymbol{X} \text{ s.t.} \left(\forall i, f_i(\mathbf{x}) \leq f_i(\mathbf{x}^*)\right) \wedge \left(\exists j, f_j(\mathbf{x}) \leq f_j(\mathbf{x}^*)\right) \tag{3}$$

The mapping of all Pareto optimal solutions in the space of objective functions is called Pareto Front and is denoted as:

$$\mathcal{P} = \left\{\left(f_1(\mathbf{x}^*), f_2(\mathbf{x}^*), \ldots, f_k(\mathbf{x}^*)\right) \in \mathbb{R}^k \,\middle|\, \mathbf{x}^* \text{ is a Pareto optimal solution}\right\} \tag{4}$$

In moral decision-making for autonomous driving, Pareto optimization methods make the system's decisions more rational and ethical by helping it balance multiple objectives (e.g., safety, efficiency, and fairness). A system can achieve an overall optimal balance by performing Pareto optimization of the balance between different objectives, finding a combination of weights that neither overly favours one objective nor completely ignores other objectives [40,41].

3.2.2 Policy gradient method: optimizing decision strategies

The policy gradient method optimizes the policy directly rather than the value function, and the method also has the potential for application in MORL. For the parameterised policy $\pi(a|s;\theta)$, the objective function $J(\theta)$ is the expected discount return:

$$J(\theta) = \mathbb{E}_{\tau \sim \pi_\theta}\left[\sum_{t=0}^{T-1} \gamma^t r_{t+1}\right] \tag{5}$$

Where $\tau =$ （$s_0, a_0, r_1, s_1, \ldots$） denotes the trajectory and $\gamma \in [0,1]$ is the discount factor. The strategy gradient theorem shows that the gradient of the objective function is:

$$\nabla_\theta J(\theta) = \mathbb{E}_{s \sim d^\pi, a \sim \pi_\theta}\left[\nabla_\theta \ln \pi(a|s;\theta) \cdot Q^\pi(s,a)\right] \tag{6}$$

Where $d^\pi(s)$ is the distribution of discounted states under the policy π_θ, representing the frequency of visits to state s weighted by the discount. $Q^\pi(s,a)$ is the state-action value function defined as the expected discounted reward for following the policy π_θ after choosing the action a at state s. In the moral decision-making problem in autonomous driving, the vehicle faces multiple possible behavioural choices, such as protecting passengers and pedestrians, obeying the law, etc. Thus, a complex reward function needs to be designed. Let the reward for the -th goal be $R_i(s,a)$, and the total reward is:

$$R(s,a) = \sum_{i=1}^{k} \omega_i R_i(s,a) \tag{7}$$

Where $\omega_i \geq 0$ is the weight, which satisfies $\sum \omega_i = 1$. For example:

① Pedestrian safety: $R_1(s,a) = -\frac{1}{d_p(s,a)+\epsilon}$ (d_p: minimum distance to pedestrians)
② Passage efficiency: $R_2(s,a) = v_{\text{avg}}(s,a)$
③ Regulatory compliance: $R_3(s,a) = -\sum \delta(v_{\text{over}}(s,a))$ (δ: overspeed function)
Substituting the composite reward into the policy gradient formula, the gradient is updated to:

$$\nabla_\theta J(\theta) = \mathbb{E}_\tau \left[\sum_{t=0}^{T-1} \nabla_\theta \log \pi_\theta(a_t|s_t) \cdot \left(\sum_{i=1}^{k} \omega_i Q_i^\pi(s_t, a_t) \right) \right] \tag{8}$$

$Q_i^\pi(s_t,\ a_t)$ is the action value function of the -th goal. To reduce variance, a baseline $b(s_t)$ is introduced:

$$\nabla_\theta J(\theta) = \mathbb{E}_\tau \left[\sum_{t=0}^{T-1} \nabla_\theta \log \pi_\theta(a_t|s_t) \cdot \left(\sum_{i=1}^{k} \omega_i \big(Q_i^\pi(s_t, a_t) - b(s_t) \big) \right) \right] \tag{9}$$

The traditional value function approach must compute each possible action's value. In contrast, the policy gradient approach needs to quantify these ethical guidelines into a reward function and find the optimal solution in the policy space by optimizing the policy parameters to achieve the best decision ultimately [42,43].

4 DISCUSSION

4.1 Weight Distribution of the Reward Function

In autonomous driving moral decision-making, the importance of different objectives may change in different contexts. In MORL, weight assignment determines how to handle and optimize multiple reward functions and directly affects the final decision-making effect of the system [44]. By assigning a weight to each objective, a composite reward function can be formed that comprehensively evaluates each objective's accomplishment. The weight assignment not only affects the system's learning process in training but also influences the moral judgement of the system in actual operation. However, in the current moral decision-making based on the MORL algorithm, the weighting problem has never been adequately solved, mainly due to:

Firstly, the choice of weights is highly subjective, and different development teams, stakeholders, and users may judge the priority of goals differently. Some developers may be more inclined to protect the safety of vehicle occupants, while others may emphasize the protection of pedestrians [45]. In addition, priority setting for moral decision-making varies across cultures, laws, and societal values. In some countries, the law may explicitly state that cars must prioritize the safety of pedestrians in all situations, whereas in others, different rules may exist [46]. Therefore, a standardized weighting scheme is unsuitable for all situations, and weighting must be flexible to accommodate different cultural and social contexts and legal environments.

Second, the environment faced by an autonomous driving system is dynamically changing, and the system must react quickly to these changes. In daytime and nighttime traffic scenarios, an autonomous driving system may need to dynamically adjust the weights for safety and efficiency [47]. In an emergency where a traffic accident occurs, the system may need to immediately increase the weighting of the safety objective, whereas, in daily traffic, more attention needs to be paid to efficiency and comfort. This context dependency requires that the weights can be dynamically adjusted according to the current state of the environment rather than being fixed. However, it is still challenging to effectively design a dynamic adjustment mechanism so that the weights can accurately reflect the needs of the current context.

Finally, in multi-objective optimization, weighting conflicts are inevitable. Safety and efficiency are often conflicting objectives in autonomous driving: higher safety usually means slower speeds and higher travelling caution, while higher efficiency may mean more risk-taking. Therefore, the system must find a suitable balance between the multiple objectives, neither sacrificing safety for efficiency nor disabling traffic flow for absolute safety. This conflict makes weight setting very fine and sensitive, and any adjustment of the weights may cause the system to deviate from one of the objectives. If the weights for safety are set too high, the system may become too conservative and even lead to traffic congestion, while if the weights for efficiency are too high, the risk of accidents may increase. Therefore, designing a reasonable weight allocation scheme to balance the weight conflicts of different objectives is a complex problem for MORL.

In today's wildly advancing autonomous driving technology, the weight distribution of the reward function is like a sword of Damocles, which has transcended the scope of mere technology and become a hub connecting algorithmic ethics and social values. It is a warning that technological development must keep pace with ethical evolution and that we realize the Rome Declaration only through verifiable technological solutions, standardized review processes, and rule-of-law-based participatory mechanisms. This technological practice is not only about the innovation of transport modes but also the reconfirmation and inheritance of humanity's ethical system in the age of AI. A multi-level weighted governance system is being constructed globally, from *ISO 21448 Standards* to UNECE regulations, MIT ethical machine experiments, and German legislative practices. More and more open, transparent, and inclusive weighted governance mechanisms are promoting autonomous driving as an intelligent carrier that carries human values.

4.2 The Universality of Ethical Objectives

The universality of ethical goals is another central challenge in moral decision-making in autonomous driving systems. Fundamental to the universality problem is ensuring that autonomous driving systems can always make moral and ethical decisions in various complex and changing social and cultural contexts.

On the one hand, morality and ethics are understood differently in different social contexts, cultural values and legal systems. In European and American cultures, the rights and freedoms of the individual are often emphasized. There may be a greater tendency to protect the lives of individuals in moral decision-making. However, in some East Asian cultures, collective interests and social harmony may be more important than individual interests [48]. Therefore, when an autonomous driving system is faced with the need to choose between protecting passengers and protecting pedestrians, different cultures may have different definitions of the ‘best choice’, and cultural differences make it challenging to devise a uniform framework for moral decision-making.

On the other hand, traffic regulations and legal systems vary significantly from country to country, which makes it difficult to harmonize moral decision-making for autonomous driving systems. In Germany, the law clearly states that innocent pedestrians must be protected as much as possible in all circumstances. At the same time, in the United States, the priorities of autonomous driving systems may focus more on the safety of passengers. In addition, different countries have inconsistent regulatory standards for autonomous driving technology. Some countries have established strict regulatory restrictions on autonomous driving behaviour, while others may give more flexibility [49]. Autonomous driving systems must be able to adapt their decisions to different legal environments to ensure the legality of their behaviour, which places a higher demand on the universality of ethical goals.

5 CONCLUSION

Although the current moral decision-making framework for autonomous driving based on MORL provides a technical basis for solving ethical dilemmas, it still faces many practical challenges. Firstly, significant subjectivity and cultural differences exist in allocating reward function weights, but the definition of ‘optimal solution’ varies significantly across societies and cultures. In addition, the fragmentation of global traffic regulations further exacerbates the difficulty of weighting standardization. Second, the problem of real-time adaptation to dynamic environments is prominent. Sudden accidents in extreme scenarios require the system to be able to adjust the weights dynamically. However, balancing algorithmic complexity and real-time decision-making with existing technologies is difficult. Finally, the scarcity of ethical dilemma data restricts model training. The probability of ‘tram dilemma’ events in real scenarios is extremely low, while the deviation of simulation data from reality may lead to insufficient decision generalization.

Future research must focus on three significant directions to address the above challenges. Firstly, a context-aware dynamic weight adjustment mechanism must be developed. The system can optimize the target weights in real time by integrating multi-dimensional parameters such as weather, road conditions, traffic density, etc. Second, to build a cross-cultural ethical framework. Establish a transparent public participation mechanism, incorporate multiple values, and coordinate regulatory conflicts through international cooperation. Third, the reality verification and credible assessment system must be strengthened. This requires relying on large-scale simulation tests and accurate road data to verify the performance of models in complex ethical dilemmas and introducing third-party auditing organizations to review the fairness and interpretability of algorithms independently. In addition, interdisciplinary collaboration (ethicists, engineers, legal experts) and public education (e.g., ethical preference surveys) will fuel the deeper integration of technology and social values.

Looking ahead, moral decision-making in autonomous driving is not only a matter of technical optimization but also of reconstructing the human ethical system in the age of intelligence. Intelligent systems are expected to achieve more humane trade-offs in extreme scenarios through dynamic weight allocation, culturally adaptive modelling, and rigorous reality verification. This process needs to be based on technological iteration and supported by global collaboration and standardized governance, ultimately enabling autonomous driving to go beyond the attribute of a tool and become a credible intelligence carrying social consensus.

COMPETING INTERESTS

The authors have no relevant financial or non-financial interests to disclose.

FUNDING

This study was supported by Shenzhen University High-level University Phase III Construction Project (000001032029) and Shenzhen University Special Funding for Ideological and Political Education (24MSZX09)

REFERENCES

[1] Chen L, Li Y, Huang C, et al. Milestones in autonomous driving and intelligent vehicles: Survey of surveys. IEEE Transactions on Intelligent Vehicles, 2022, 8(2): 1046–1056.

[2] Geisslinger M, Poszler F, Betz J, et al. Autonomous driving ethics: from trolley problem to ethics of risk. Philosophy & Technology, 2021, 34(4): 1033–1055.
[3] Liu Q, Li X, Yuan S, et al. Decision-making technology for autonomous vehicles: Learning-based methods, applications and future outlook. 2021 IEEE International Intelligent Transportation Systems Conference (ITSC), IEEE, 2021: 30–37.
[4] Wang H, Khajepour A, Cao D, et al. Ethical decision making in autonomous vehicles: challenges and research progress. IEEE Intelligent Transportation Systems Magazine, 2020, 14(1): 6–17.
[5] Li W, Devidze R, Mustafa W, et al. Ethics in Action: Training Reinforcement Learning Agents for Moral Decision-making in Text-based Adventure Games. International Conference on Artificial Intelligence and Statistics, 2024: 1954–1962.
[6] Ecoffet A, Lehman J. Reinforcement learning under moral uncertainty. International Conference on Machine Learning, 2021: 2926–2936.
[7] Noothigattu R, Bouneffouf D, Mattei N, et al. Teaching AI agents ethical values using reinforcement learning and policy orchestration. IBM Journal of Research and Development, 2019, 63(4/5): 2–1.
[8] Abel D, MacGlashan J, Littman ML. Reinforcement learning as a framework for ethical decision making. Workshops at the Thirtieth AAAI Conference on Artificial Intelligence, 2016.
[9] Deschamps T, Chaput R, Matignon L. Multi-objective reinforcement learning: an ethical perspective. RJCIA, 2024.
[10] Tennant E, Hailes S, Musolesi M. Modelling moral choices in social dilemmas with multi-agent reinforcement learning. Proceedings of the Thirty-Second International Joint Conference on Artificial Intelligence, IJCAI-23, 2023: 317–325.
[11] Rodriguez-Soto M, Serramia M, Lopez-Sanchez M, et al. Instilling moral value alignment by means of multi-objective reinforcement learning. Ethics and Information Technology, 2022, 24(1): 9.
[12] Peschl M, Zgonnikov A, Oliehoek FA, et al. MORAL: Aligning AI with human norms through multi-objective reinforced active learning. arXiv preprint arXiv:2201.00012, 2021.
[13] Bruers S, Braeckman J. A review and systematization of the trolley problem. Philosophia, 2014, 42: 251–269.
[14] Thomson JJ. The trolley problem. Yale LJ, 1984, 94: 1395.
[15] Thomson JJ. Killing, letting die, and the trolley problem. The Monist, 1976: 204–217.
[16] Paulo N. The Trolley Problem in the Ethics of Autonomous Vehicles. The Philosophical Quarterly, 2023, 73(4): 1046–1066.
[17] Wu SS. Autonomous vehicles, trolley problems, and the law. Ethics and Information Technology, 2020, 22(1): 1–13.
[18] Wiseman Y, Grinberg I. The Trolley Problem Version of Autonomous Vehicles. The Open Transportation Journal, 2018, 12(1).
[19] Himmelreich J. Never mind the trolley: The ethics of autonomous vehicles in mundane situations. Ethical Theory and Moral Practice, 2018, 21(3): 669–684.
[20] Moolayil AK. The modern trolley problem: ethical and economically sound liability schemes for autonomous vehicles. Case W. Res. JL Tech. & Internet, 2018, 9: 1.
[21] Etienne H. When AI ethics goes astray: a case study of autonomous vehicles. Social Science Computer Review, 2022, 40(1): 236–246.
[22] Tolmeijer S, Christen M, Kandul S, et al. Comparing AI and human expert collaboration in ethical decision making. Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems, 2022: 1–17.
[23] Madhav AS, Tyagi AK. Explainable Artificial Intelligence (XAI): Connecting artificial decision-making and human trust in autonomous vehicles. Proceedings of Third International Conference on Computing, Communications, and Cyber-Security: IC4S 2021, Springer Nature Singapore, 2022: 123–136.
[24] Martinho A, Herber N, Kroesen M, et al. Ethical issues in focus by the autonomous vehicles industry. Transport Reviews, 2021, 41(5): 556–577.
[25] Nassar A, Kamal M. Ethical dilemmas in AI-powered decision-making: a deep dive into big data-driven ethical considerations. International Journal of Responsible Artificial Intelligence, 2021, 11(8): 1–11.
[26] Conitzer V, Sinnott-Armstrong W, Borg JS, et al. Moral decision-making frameworks for artificial intelligence. Proceedings of the AAAI Conference on Artificial Intelligence, 2017, 31(1).
[27] Noothigattu R, Gaikwad S, Awad E, et al. A voting-based system for ethical decision-making. Proceedings of the AAAI Conference on Artificial Intelligence, 2018, 32(1).
[28] Qian L, Yin J, Huang Y, et al. The role of values and ethics in influencing consumers' intention to use autonomous vehicle hailing services. Technological Forecasting and Social Change, 2023, 188: 122267.
[29] Ernst D, Louette A. Introduction to reinforcement learning. Feuerriegel S, Hartmann J, Janiesch C, et al., eds. AI and Data Science in Business Applications, 2024: 111–126.
[30] Wiering MA, Van Otterlo M. Reinforcement learning. Adaptation, Learning, and Optimization, 2012, 12(3): 729.
[31] Sutton R S, Barto A G. Reinforcement learning. Journal of Cognitive Neuroscience, 1999, 11(1): 126–134.
[32] Barto A G. Reinforcement learning. In Neural systems for control (pp. 7–30). Academic Press, 1997.
[33] Kaelbling L P, Littman M L, Moore A W. Reinforcement learning: a survey. Journal of Artificial Intelligence Research, 1996, 4: 237–285.

[34] Kiran B R, Sobh I, Talpaert V, et al. Deep reinforcement learning for autonomous driving: a survey. IEEE Transactions on Intelligent Transportation Systems, 2021, 23(6): 4909–4926.
[35] Kendall A, Hawke J, Janz D, et al. Learning to drive in a day. In 2019 International Conference on Robotics and Automation (ICRA) (pp. 8248–8254). IEEE, 2019.
[36] Singh J. The Ethics of Data Ownership in Autonomous Driving: Navigating Legal, Privacy, and Decision-Making Challenges in a Fully Automated Transport System. Australian Journal of Machine Learning Research & Applications, 2022, 2(1): 324–366.
[37] Titu A M, Stanciu A, Mihaescu L. Technological and ethical aspects of autonomous driving in a multicultural society. In 2020 12th International Conference on Electronics, Computers and Artificial Intelligence (ECAI) (pp. 1–6). IEEE, 2020.
[38] Wang Z, Rangaiah G P. Application and analysis of methods for selecting an optimal solution from the Pareto-optimal front obtained by multiobjective optimization. Industrial & Engineering Chemistry Research, 2017, 56(2): 560–574.
[39] Deb K, Gupta H. Searching for robust Pareto-optimal solutions in multi-objective optimization. In International Conference on Evolutionary Multi-Criterion Optimization (pp. 150–164). Berlin, Heidelberg: Springer Berlin Heidelberg, 2005.
[40] Yan Y, Du H, He D, Li W. Pareto optimal information flow topology for control of connected autonomous vehicles. IEEE Transactions on Intelligent Vehicles, 2022, 8(1): 330–343.
[41] MacLeod B P, Parlane F G, Rupnow C C, et al. A self-driving laboratory advances the Pareto front for material properties. Nature Communications, 2022, 13(1): 995.
[42] Li G, Li S, Qu X. Continuous decision-making for autonomous driving at intersections using deep deterministic policy gradient. IET Intelligent Transport Systems, 2022, 16(12): 1669–1681.
[43] Hu H, Lu Z, Wang Q, Zheng C. End-to-end automated lane-change manoeuvring considering driving style using a deep deterministic policy gradient algorithm. Sensors, 2020, 20(18): 5443.
[44] Hayes C F, Rădulescu R, Bargiacchi E, et al. A practical guide to multi-objective reinforcement learning and planning. Autonomous Agents and Multi-Agent Systems, 2022, 36(1): 26.
[45] Millard-Ball A. Pedestrians, autonomous vehicles, and cities. Journal of Planning Education and Research, 2018, 38(1): 6–12.
[46] World Health Organization. Pedestrian safety: a road safety manual for decision-makers and practitioners. World Health Organization, 2023.
[47] Muhammad K, Ullah A, Lloret J, et al. Deep learning for safe autonomous driving: current challenges and future directions. IEEE Transactions on Intelligent Transportation Systems, 2020, 22(7): 4316–4336.
[48] Ho D. Asian vs. liberal democracy: Identifying the locus of conflict in the Asian values debate. Political Science, 2023, 75(2): 165–192.
[49] Taeihagh A, Lim H S M. Governing autonomous vehicles: emerging responses for safety, liability, privacy, cybersecurity, and industry risks. Transport Reviews, 2019, 39(1): 103–128.